





**CLOVER MS Data Analysis Software** 

1. A list of the main features that co in the classification: the higher the the more important it is to discrimin

2. A graph in which every sample each feature. This plot gives us diffe

2.1. The more to the right the point contribution the point has in the the point is on the left side, its opposite category.

2.2. The colour indicates if the feat or a low value (in blue).

In this example, the first five featu 938) have a high contribution picked category. Indeed, these positive class identification.

<complex-block>2 And the set of the</complex-block>	
Intaining Clover BioSoft © Logout     Intaining Clover BioSoft     I	
Options :       Image: Classification         Image: Classification       Image: Classification	ng Clove cation th spectru lassifier ifferentiat plot and e similarit /are st Class
Outlier Detection       Repository       Classification       A category will be assigned to each spectrum on the peak matrix. Then, the KNN classifier builds a model which try to differentiate those categories.       A nalysis       Analysis       Test       A lignmen         V1.6.10 (#87e52aa1)       N classification       I bio Analysis       I classification       I classificati	pital 1 pital 2 roducibilit ent eriments v
Conception of the second seco	Reproducib will be us   2 × 9b0e-e2bd 3-bc79-42f I-a573-168 d-a019-3d; -b31f-b11(
Create a Peak Matrix 3. Activate the button of Shapley values and train the model. 4. Create a Peak Matrix 4. Create a	•
ontribute (to have a function) e feature position in the list is, inate your picked category.	
is represented by a point for ferent information:	
int is in the plot, the highest picked category. However, if function contributes to the	
ture has a high value (in red)	
ures (934, 935, 933, 937, and in the discrimination of the features contribute to the	Al-p accu Resi

### CLOVER BioSoft

### 1. Open the Classification section and click to perform Random Forest analysis.

🛛 🖉 Training Clover [training@cloverbiosoft.com] 🛛 Premium 🏛 Training Clover BioSoft 🛛 🔂 Logo

Search... Q

118 files 🗋

113 files 🗋

231 files 🗋

Samples: 184

nina@cloverbiosoft.com | 23-10-2023 13:12:15 GMT+2 🗹 Alignment 🗹 Merge Normalization mber of experiments visualized: 3 🙎 Training Clover [training@cloverbiosoft.com] 🛛 Premium 🏛 Training Clover BioSoft 🛛 🔂 Logo m Forest Classifier Pneumoniae Assay Experiment Reprodu e the categories and samples that will be used as inputs for the analysis. You can remove, add or create more categories, and move 🗙 Clear 🛛 🔂 Add all category Hospital 1 🗙 💦 Hospital 2 🗙 🖍 Edit Category 🔄 niae\_Kp1\_A1\_4a0783cf-d2f5-4fcc-

oniae Kp5 E1 c8d5e939-4de0-4ed3-bc79-42fd45d1720 noniae\_Kp5\_E5\_75f95aea-21a8-4e0d-a573-1688893c374c K. Pneumoniae/Hospital 1/K\_pneumoniae\_Kp5\_E5\_7 oniae\_Kp5\_E9\_31213102-31c3-423d-a019-3d318468c4c0 toniae\_Kp6\_F1\_c54a5556-47fb-47bf-b31f-b110423e4147 K. Pneumoniae/Hospital 1/K\_pneumoniae\_Kp6\_F1\_c54a5556-47fb-47bf-b31f-b110423e4147.0.xml.bkp

n] 🛛 Premium 🏛 Training Clover BioSoft 🛛 🔂 Logou

ndom Forest Classification - Select Peak Matrix

ase, select an experiment and then a peak matrix:

Reproducibility between Hospitals





cloverbiosoft.com clovermsdataanalysis.com

**Clover Bioanalytical Software** Copyright<sup>®</sup> 2015

Edif. Centro de Empresas PTS. Av. del Conocimiento, 41 18016 Granada, Spain

Phone: +34 958 991 543 email: info@cloverbiosoft.com support@cloverbiosoft.com

Clover MS Data Analysis Software, Quick Start Guide

Al-powered software for research on fast, accurate and cost-efficient Microorganisms Resistance Analysis and Strain Typing using both MALDI-MS and FTIR data

HOME SOFTWARE - TESTIMONIALS TIMELINE ABOUT US MANUAL SIGN UP LOGIN CONTACT US

1876.35615 793.40392 651.73132 2461.43972 2502.51022 1279.56051 825.36207 1782.18165 751.92235 694.81792 1848.84887 515.54349 2176.8864 1265.51052 1702.80763 923.01272 638.3717 1966.66387 1862.37375 881.03365

## How to perform Shapley values in our platform?

**CLOVER** MS Data Analysis Software





Mean abs(shap value)

Shap value

In Clover Biosoft, it is possible to perform Shapley values in a Random Forest analysis. In our case, Shapley values measure the feature importance, based on its average marginal contribution considering all possible combinations, in order to establish a classification between different categories.

# CLOVER BioSoft

![](_page_1_Picture_12.jpeg)

cloverbiosoft.com clovermsdataanalysis.com

**Clover Bioanalytical Software** Copyright<sup>®</sup> 2015

HOME SOFTWARE - TESTIMONIALS TIMELINE ABOUT US MANUAL SIGN UP LOGIN CONTACT US

Al-powered software for research on fast, accurate and cost-efficient Microorganisms Resistance Analysis and Strain Typing using both MALDI-MS and FTIR data

CLOVER BioSoft

Edif. Centro de Empresas PTS. Av. del Conocimiento, 41 18016 Granada, Spain

Phone: +34 958 991 543 email: info@cloverbiosoft.com support@cloverbiosoft.com

![](_page_2_Picture_0.jpeg)

**CLOVER MS Data Analysis Software** 

CLOVER BioSoft

# CLOVER BioSoft

![](_page_2_Picture_9.jpeg)

cloverbiosoft.com clovermsdataanalysis.com

**Clover Bioanalytical Software** Copyright<sup>®</sup> 2015

HOME SOFTWARE - TESTIMONIALS TIMELINE ABOUT US MANUAL SIGN UP LOGIN CONTACT US

Al-powered software for research on fast, accurate and cost-efficient Microorganisms Resistance Analysis and Strain Typing using both MALDI-MS and FTIR data

Edif. Centro de Empresas PTS. Av. del Conocimiento, 41 18016 Granada, Spain

Phone: +34 958 991 543 email: info@cloverbiosoft.com <u>support@cloverbiosoft.co</u>m

![](_page_3_Picture_0.jpeg)

**CLOVER** MS Data Analysis Software

![](_page_3_Picture_6.jpeg)

1. Open the Classification section and click to perform Random Forest analysis.

	은 Traini	ng Clover [training@clov	verbiosoft.com	] Premium 🏛 T	Fraining Clover BioSof	t 🕞 Logout
on - Selec	t Peak Matrix					
peak matrix.						
					Search	0
					Search	118 files <b>[</b> ]
						113 files 🖪
Hospitals	2					231 files 🕒
Merge	✓ Normalizatio	on	t	raining@cloverbiosof	t.com   23-10-2023 13:	12:15 GMT+2 Samples: 184
3/3						
8	Training Clover [training@cl	overbiosoft.com] Prem	ium 🏛 Trai	ning Clover BioSoft	🕞 Logout	
Hospitals	Peak Matrix Test					
s for the anal	ysis. You can remove, add o	r create more categories,	and move	× Clear	• Add	
				an	category	
				×	Edit Category	
K. Pneumoniae	/Hospital 1/K_pneumoniae_Kp1	_A1_4a0783cf-d2f5-4fcc-9b0	)e-e2bd73237ee	d.0.xml.bkp		
K. Pneumonia	e/Hospital 1/K_pneumoniae_Kp	5_E1_c8d5e939-4de0-4ed3-b	c79-42fd45d172	205.0.xml.bkp		
K. Pneumonia	ae/Hospital 1/K_pneumoniae_Kp	05_E9_31213102-31c3-423d-	a019-3d318468	c4c0.0.xml.bkp		
K. Pneumoniae	/Hospital 1/K_pneumoniae_Kp6	_F1_c54a5556-47fb-47bf-b3	1f-b110423e414	7.0.xml.bkp		
<b>**</b>	1 2 3 4 5 19 » x	»»»				
defeul N	Max depth	Min samples split	Min sam	oles per leaf		
	None - dela 🗸	2 - default 🗸	I - den	auit 🗸		
3						
lues	Train Random Forest mode					

CLOVER BioSoft

# CLOVER BioSoft

![](_page_3_Picture_11.jpeg)

cloverbiosoft.com clovermsdataanalysis.com

**Clover Bioanalytical Software** Copyright<sup>®</sup> 2015

HOME SOFTWARE - TESTIMONIALS TIMELINE ABOUT US MANUAL SIGN UP LOGIN CONTACT US

Al-powered software for research on fast, accurate and cost-efficient Microorganisms Resistance Analysis and Strain Typing using both MALDI-MS and FTIR data

Edif. Centro de Empresas PTS. Av. del Conocimiento, 41 18016 Granada, Spain

Phone: +34 958 991 543 email: info@cloverbiosoft.com <u>support@cloverbiosoft.co</u>m

![](_page_4_Figure_0.jpeg)

**CLOVER** MS Data Analysis Software

1. A list of the main features that contribute (to have a function) in the classification: the higher the feature position in the list is, the more important it is to discriminate your picked category.

2. A graph in which every sample is represented by a point for each feature. This plot gives us different information:

2.1. The more to the right the point is in the plot, the highest contribution the point has in the picked category. However, if the point is on the left side, its function contributes to the opposite category.

2.2. The colour indicates if the feature has a high value (in red) or a low value (in blue).

In this example, the first five features (934, 935, 933, 937, and 938) have a high contribution in the discrimination of the picked category. Indeed, these features contribute to the positive class identification.

### CLOVER BioSoft

![](_page_4_Picture_9.jpeg)

cloverbiosoft.com clovermsdataanalysis.com

Clover Bioanalytical Software Copyright® 2015

HOME SOFTWARE - TESTIMONIALS TIMELINE ABOUT US MANUAL SIGN UP LOGIN CONTACT US

Al-powered software for research on fast, accurate and cost-efficient Microorganisms Resistance Analysis and Strain Typing using both MALDI-MS and FTIR data

Edif. Centro de Empresas PTS. Av. del Conocimiento, 41 18016 Granada, Spain

Phone: +34 958 991 543 email: info@cloverbiosoft.com <u>support@cloverbiosoft.co</u>m